PATENT

IN THE TITLE

Please amend the title per the Examiner's suggestion to read:

"METHOD AND APPARATUS FOR PREDICTIVELY QUANTIZING VOICED SPEECH WITH SUBTRACTION OF WEIGHTED PARAMETERS FOR PREVIOUS FRAMES"

IN THE SPECIFICATION

Applicants provide herewith amendments to the specification. A marked up version of these amendments is also provided on a separate page from this Amendment as Appendix A.

Please amend the paragraphs of the specification as follows:

On page 5, beginning at line 10, paragraph no. 1:

A1

Time-domain coders such as the CELP coder typically rely upon a high number of bits, N_0, per frame to preserve the accuracy of the time-domain speech waveform. Such coders typically deliver excellent voice quality provided the number of bits, N_0, per frame is relatively large (e.g., 8 kbps or above). However, at low bit rates (4 kbps and below), time-domain coders fail to retain high quality and robust performance due to the limited number of available bits. At low bit rates, the limited codebook space clips the waveform-matching capability of conventional time-domain coders, which are so successfully deployed in higher-rate commercial applications. Hence, despite improvements over time, many CELP coding systems operating at low bit rates suffer from perceptually significant distortion typically characterized as noise.

On page 7, beginning at line 13, paragraph no. 3:

In recent years, coders have emerged that are hybrids of both waveform coders and parametric coders. Illustrative of these so-called hybrid coders is the prototype-waveform interpolation (PWI) speech coding system. The PWI coding system may also be known as a prototype pitch period (PPP) speech coder. A PWI coding system provides an efficient method for coding voiced speech. The basic concept of PWI is to extract a representative pitch cycle (the prototype waveform) at fixed intervals, to transmit its description, and to reconstruct the speech signal by interpolating between the prototype waveforms. The PWI method may operate either on the LP residual signal or on the speech signal. An exemplary PWI, or PPP, speech coder is described in U.S. Application Serial No. 09/217,494, entitled PERIODIC SPEECH CODING, filed December 21, 1998, now U.S. Patent No. 6,456,964 issued October 24, 2002, assigned to the assignee of the present invention, and fully incorporated herein by reference. Other PWI, or PPP, speech coders are described in U.S. Patent No. 5,884,253 and W. Bastiaan Kleijn & Wolfgang Granzow *Methods for Waveform Interpolation in Speech Coding, in* 1 *Digital Signal Processing* 215-230 (1991).

A2

On page 13, beginning at line 8, paragraph no. 1:

During typical operation of the cellular telephone system, the base stations 12 receive sets of reverse link signals from sets of mobile units 10. The mobile units 10 are conducting telephone calls or other communications. Each reverse link signal received by a given base station 12 is processed within that base station 12. The resulting data is forwarded to the BSC 14. The BSC 14 provides call resource allocation and mobility management functionality including the orchestration of soft

A3

handoffs between base stations 12. The BSC 14 also routes the received data to the MSC 16, which provides additional routing services for interface with the PSTN 18. Similarly, the PSTN 18 interfaces with the MSC 16, and the MSC 16 interfaces with the BSC 14, which in turn control the base stations 12 to transmit sets of forward link signals to sets of mobile units 10. It should be understood by those of skill that the subscriber units 10 may be fixed units in alternate embodiments.

On page 14, beginning at line 5, paragraph no. 1:

A4

The speech samples s(n) represent speech signals that have been digitized and quantized in accordance with any of various methods known in the art including, e.g., pulse code modulation (PCM), companded μ-law, or A-law. As known in the art, the speech samples s(n) are organized into frames of input data wherein each frame comprises a predetermined number of digitized speech samples s(n). In an exemplary embodiment, a sampling rate of 8 kHz is employed, with each 20 ms frame comprising 160 samples. In the embodiments described below, the rate of data transmission may advantageously be varied on a frame-by-frame basis from full rate to half rate to quarter rate to eighth rate. Varying the data transmission rate is advantageous because lower bit rates may be selectively employed for frames containing relatively less speech information. As understood by those skilled in the art, other sampling rates and/or frame sizes may be used. Also in the embodiments described below, the speech encoding (or coding) mode may be varied on a frame-by-frame basis in response to the speech information or energy of the frame.

On page 14, beginning at line 20, paragraph no. 2:

A5

The first encoder 100 and the second decoder 110 together comprise a first speech coder (encoder/decoder), or speech codec. The speech coder could be used in any communication device for transmitting speech signals, including, e.g., the subscriber units, BTSs, or BSCs described above with reference to FIG. 1. Similarly, the second encoder 106 and the first decoder 104 together comprise a second speech coder. It is understood by those of skill in the art that speech coders may be

implemented with a digital signal processor (DSP), an application-specific integrated circuit (ASIC), discrete gate logic, firmware, or any conventional programmable software module and a microprocessor. The software module could reside in RAM memory, flash memory, registers, or any other form of storage medium known in the art. Alternatively, any conventional processor, controller, or state machine could be substituted for the microprocessor. Exemplary ASICs designed specifically for speech coding are described in U.S. Patent No. 5,727,123, assigned to the assignee of the present invention and fully incorporated herein by reference, and U.S. Application Serial No. 08/197,417, entitled VOCODER ASIC, filed February 16, 1994, now U.S. Patent No. 5,784,532 issued July 21, 1998, assigned to the assignee of the present invention, and fully incorporated herein by reference.

A5

On page 17, beginning at line 12, paragraph no. 2:

In one embodiment, illustrated in FIG. 5, a multimode speech encoder 400 communicates with a multimode speech decoder 402 across a communication channel, or transmission medium, 404. The communication channel 404 is advantageously an RF interface configured in accordance with the IS-95 standard. It would be understood by those of skill in the art that the encoder 400 has an associated decoder (not shown). The encoder 400 and its associated decoder together form a first speech coder. It would also be understood by those of skill in the art that the decoder 402 has an associated encoder (not shown). The decoder 402 and its associated encoder together form a second speech coder. The first and second speech coders may advantageously be implemented as part of first and second DSPs, and may reside in, e.g., a subscriber unit and a base station in a PCS or cellular telephone system, or in a subscriber unit and a gateway in a satellite system.

A6

On page 18, beginning at line 13, paragraph no. 2:

A speech signal, *s(n)*, is provided to the parameter calculator 406. The speech signal is divided into blocks of samples called frames. The value *n* designates the frame number. In an alternate embodiment, a linear prediction (LP) residual error signal is used in place of the speech signal. The LP residue is used by speech coders such as, e.g., the CELP coder. Computation of the LP residue is advantageously performed by providing the speech signal to an inverse LP filter (not shown). The transfer function of the inverse LP filter, *A(z)*, is computed in accordance with the following equation:

A7

$$A(z) = 1 - a_1 z^{-1} - a_2 z^{-2} - \ldots -a_p z^{-p},$$

in which the coefficients a_l are filter taps having predefined values chosen in accordance with known methods, as described in the aforementioned U.S. Patent No. 5,414,796 and U.S. Patent No. 6,456,964. The number *p* indicates the number of previous samples the inverse LP filter uses for prediction purposes. In a particular embodiment, *p* is set to ten.

On page 22, beginning at line 1, paragraph no. 1:

In accordance with a NELP encoding mode 410, a filtered, pseudo-random noise signal is used to model the speech frame. The NELP encoding mode 410 is a relatively simple technique that achieves a low bit rate. The NELP encoding mode 410 may be used to advantage to encode frames classified as unvoiced speech. An exemplary NELP encoding mode is described in detail in the aforementioned U.S. Patent No. 6,456,964.

A8

On page 22, beginning at line 7, paragraph no. 2:

In accordance with a PPP encoding mode 410, only a subset of the pitch periods within each frame are encoded. The remaining periods of the speech signal are reconstructed by interpolating between these prototype periods. In a time-domain implementation of PPP coding, a first set of parameters is calculated that describes how to modify a previous prototype period to approximate the current prototype period. One or more codevectors are selected which, when summed, approximate the difference between the current prototype period and the modified previous prototype period. A second set of parameters describes these selected codevectors. In a frequency-domain implementation of PPP coding, a set of parameters is calculated to describe amplitude and phase spectra of the prototype. This may be done either in an absolute sense, or predictively as described hereinbelow. In either implementation of PPP coding, the decoder synthesizes an output speech signal by reconstructing a current prototype based upon the first and second sets of parameters. The speech signal is then interpolated over the region between the current reconstructed prototype period and a previous reconstructed prototype period. The prototype is thus a portion of the current frame that will be linearly interpolated with prototypes from previous frames that were similarly positioned within the frame in order to reconstruct the speech signal or the LP residual signal at the decoder (i.e., a past prototype period is used as a predictor of the current prototype period). An exemplary PPP speech coder is described in detail in the aforementioned U.S. Patent No. 6,456,964.

On page 24, beginning at line 9, paragraph no. 2:

If the packet disassembler and packet loss detector module 414 detects the packet, the packet is disassembled and provided to the pertinent decoding mode 416. If the packet disassembler and packet loss detector module 414 does not detect a packet, a packet loss is declared and the erasure decoder 418 advantageously performs frame erasure processing as described in a related Application No. 09/557,283 filed April 24, 2000, entitled FRAME ERASURE COMPENSATION METHOD IN A VARIABLE RATE SPEECH CODER, assigned to the assignee of the present invention, and fully incorporated herein by reference.

A10

On page 24, beginning at line 17, paragraph no. 3:

The parallel array of decoding modes 416 and the erasure decoder 418 are coupled to the post filter 420. The pertinent decoding mode 416 decodes, or de-quantizes, the packet provides the information to the post filter 420. The post filter 420 reconstructs, or synthesizes, the speech frame, outputting synthesized speech frames, $\hat{s}(n)$. Exemplary decoding modes and post filters are described in detail in the aforementioned U.S. Patent No. 5,414,796 and U.S. Patent No. 6,456,964.

A11

On page 27, beginning at line 19, paragraph no. 3:

Due to their periodic nature, voiced frames can be coded using a scheme in which the entire set of bits is used to quantize one prototype pitch period, or a finite set of prototype pitch periods, of the frame of a known length. This length of the prototype pitch period is called the pitch lag. These prototype pitch periods, and possibly the prototype pitch periods of adjacent frames, may then be used to reconstruct the entire speech frame without loss of perceptual quality. This PPP scheme of extracting the prototype pitch period from a frame of speech and using these prototypes for

A12

A12 reconstructing the entire frame is described in the aforementioned U.S. Patent No. 6,456,964.

On page 28, beginning at line 5, paragraph no. 1:

A13

In one embodiment a quantizer 500 is used to quantize highly periodic frames such as voiced frames in accordance with a PPP coding scheme, as shown in FIG. 7. The quantizer 500 includes a prototype extractor 502, a frequency domain converter 504, an amplitude quantizer 506, and a phase quantizer 508. The prototype extractor 502 is coupled to the frequency domain converter 504. The frequency domain converter 504 is coupled to the amplitude quantizer 506 and to the phase quantizer 508.

On page 29, beginning at line 1, paragraph no. 1:

A14

Other schemes for coding voiced frames, such as, e.g., multiband excitation (MBE) speech coding and harmonic coding, transform the entire frame (either LP residue or speech) or parts thereof into frequency-domain values through Fourier transform representations comprising amplitudes and phases that can be quantized and used for synthesis into speech at the decoder (not shown). To use the quantizer of FIG. 7 with such coding schemes, the prototype extractor 502 is omitted, and the frequency domain converter 504 serves to decompose the complex short-term frequency spectral representations of the frame into an amplitude vector and a phase vector. And in either coding scheme, a suitable windowing function such as, e.g., a Hamming window, may first be applied. An exemplary MBE speech coding scheme is

described in D.W. Griffin & J.S. Lim, "Multiband Excitation Vocoder," 36(8) *IEE Trans. on ASSP* (Aug. 1988). An exemplary harmonic speech coding scheme is described in L.B. Almeida & J.M. Tribolet, "Harmonic Coding: A Low Bit-Rate, Good Quality, Speech Coding Technique," *Proc. ICASSP '82* 1664-1667 (1982).

A14

On page 32, beginning at line 16, paragraph no. 2:

In one embodiment phase values may be quantized as follows. A subset of the phase vector for frame '*m*' may be denoted ϕ_m. It is possible to quantize ϕ_m as being equal to the phase of a reference waveform (time domain or frequency domain of the entire frame or a part thereof), and zero or more linear shifts applied to one or more bands of the transformation of the reference waveform. Such a quantization technique is described in U.S. Application Serial No. 09/356,491 entitled METHOD AND APPARATUS FOR SUBSAMPLING PHASE SPECTRUM INFORMATION, filed July 19, 1999, now U.S. Patent No. 6,397,175 issued May 28, 2002, assigned to the assignee of the present invention, and fully incorporated herein by reference. Such a reference waveform could be a transformation of the waveform of frame m_N, or any other predetermined waveform.

A15